

Workshop on the Application of Genomic Tools for
the Rapid Molecular Characterization of Bacterial
Isolates in Food-borne Disease Outbreak
Investigations Ottawa, ON, February 24-25, 2014

Prepared by:
Burton Blais
Canadian Food Inspection Agency

Scientific Authority:
Norman Yanofsky
DRDC Centre for Security Science

The scientific or technical validity of this Contract Report is entirely the responsibility of the Contractor and the contents do not necessarily have the approval or endorsement of the Department of National Defence of Canada.

DRDC-RDDC-2014-C144

May 2014

Canadian Food Inspection Agency

WORKSHOP REPORT

Workshop on the Application of Genomic Tools for the Rapid Molecular Characterization of Bacterial Isolates in Food-borne Disease Outbreak Investigations

Ottawa, ON, February 24-25, 2014

This report was prepared by
The Workshop Organizing Committee,
In collaboration with:



205 Catherine, Suite 300
Ottawa, Ontario K2P 1C3
Ph : 613.230.6424
Fax : 613.567.1504
www.intersol.ca

IMPORTANT INFORMATIVE STATEMENTS

CSSP-2013-CD-1056 Workshop on the application of genomics tools for the rapid molecular characterization of bacterial isolates in foodborne disease outbreak investigations was supported by the Canadian Safety and Security Program (CSSP) which is led by Defence Research and Development Canada's Centre for Security Science, in partnership with Public Safety Canada. The project was led by Canadian Food Inspection Agency with support from partners at Health Canada, Public Health Agency of Canada, and US Food and Drug Agency.

CSSP is a federally-funded program to strengthen Canada's ability to anticipate, prevent/mitigate, prepare for, respond to, and recover from natural disasters, serious accidents, crime and terrorism through the convergence of science and technology with policy, operations and intelligence

© Her Majesty the Queen in Right of Canada, as represented by the Minister of National Defence, 2014

© Sa Majesté la Reine (en droit du Canada), telle que représentée par le ministre de la Défense nationale, 2014

Revision Date: April 10, 2014

Table of Contents

Background.....	4
Setting the Context	5
Opening Remarks.....	5
Current State – Characterization of Microbial Hazards and Regulatory Decision Making.....	5
Public Health Agency of Canada: Food-borne Illness Outbreak Investigations from An Epidemiologic Perspective	5
Food-borne Disease Real Time Surveillance and Outbreak Response in Canada	6
Local and Provincial Food-borne Illness Outbreak Identification and Response	6
Application of the Weight of Evidence in Responsive Health Risk Assessments during Food-borne Outbreaks	6
Discussion	7
Current Tools for Detection, Typing, Characterization, Risk Profiling – High Risk Pathogens.....	8
Implementation of Molecular Techniques in Regulatory Food Microbiology Testing	8
From A,C,G,T to PFGE, to MLST, to MLVA, to WGS: How to Make Food Safer	8
Use of Molecular Markers for Characterization of Food-Borne Pathogens	9
Front Line Applications of New Typing Methodologies to Investigate Food-Borne Disease: Promises and Challenges.....	9
Discussion	9
New Possibilities for Genomics	11
One Disrupting Technology Fits it All: Towards Standardized Bacterial Whole Genome Sequencing for Global Surveillance	11
Discussion	13
The Path Forward.....	14
Discussion on Other Challenges and Additional Research Needed.....	14
Proposed Next Steps.....	15
Closing Remarks.....	16
Appendix A - Concept paper.....	17
Appendix B - Workshop survey results.....	26
Appendix C - Meeting Agenda.....	32
Appendix D - Participant's list.....	35
Appendix E - Workshop Organizing Committee.....	38
Appendix F - Acknowledgements.....	38

Background

Recent developments in genomics technologies are creating a major shift in the characterization and management of microbial hazards in the food supply. Increasingly, regulatory food safety specialists charged with investigating food contamination and illness outbreak events must synthesize new types of information to arrive at a final decision in determining the best course of action mitigating public health impacts.

On February 24 and 25, 2014, the Canadian Food Inspection Agency (CFIA) hosted a facilitated workshop in Ottawa, ON to focus on the application of genomic tools for the rapid molecular characterization of bacterial isolates in food-borne disease outbreak investigations. The two-day workshop brought together approximately 85 participants, including scientists and regulatory food safety experts involved in conducting food-borne disease outbreak investigations at the provincial and federal levels. The meeting agenda and list of participants can respectively be found in Appendices A and B of this report.

The workshop featured a panel of food safety experts from Canada and the international community presenting an overview of genomics technologies as applied in the identification, characterization and typing of bacterial food-borne isolates, as well as some case studies highlighting the role of genomics in the resolution of critical regulatory food microbiology issues in their respective jurisdictions.

The objectives of the workshop were:

- to explore the application of genomic tools supporting the rapid molecular characterization of bacterial isolates in food-borne disease outbreak investigations;
- to clarify the potential and limitations of pathogen genomic tools for regulatory end-users; and
- to identify research gaps that may need to be explored in order to meet regulatory needs.

The workshop provided participants with the opportunity to develop a clearer understanding of the promise and limitations of genomics technologies, and how best to integrate analytical information in their regulatory decision making processes.

What follows here is a report from this workshop. During the two-day event, several presentations were made to provide a starting point for the discussions; these presentations are only summarized briefly here as they are available under separate cover. The report includes a synthesis of key messages recorded during discussions, and is intended as a record of the meeting, to be used by the CFIA and its partners in pursuing their work in this area.

Setting the Context

Opening Remarks

Burton Blais (Workshop Chair, CFIA) welcomed participants to the workshop and provided an overview of the agenda for the two-day event. Three thematic sessions – each including a panel of speakers and interactive discussions – formed the basis of the workshop: 1) how decisions makers approach incidents and consider information to arrive at a regulatory decision; 2) some of the current approaches and tools in use for molecular characterization; and 3) developments leading to exciting new possibilities for the future. The importance of ensuring that both scientists and regulators understand the promises and limitations of genomic tools was highlighted, as this understanding will contribute to determining how best to integrate these tools to enhance food safety. The workshop was designed to help participants begin to explore these questions.

Martine Dubuc (Chief Food Safety Officer for Canada, CFIA) thanked the Canadian Safety and Security Program of Defence Research Canada for the funding provided in support of this workshop. She then shared a vision of genomics for the CFIA. As food is distributed globally, virus and bacterial contamination know no frontier. In order to enable international collaboration and sharing of data and information globally in times of crisis, standardization is critical. As science evolves and new tools become available, it is important to consider what each new piece of information translates into as it relates to addressing food safety challenges. The possibility of developing tools that the industry could use in order to improve the ability to ensure the safety of foods represents another question for consideration.

She also indicated that Canada is among nine countries taking part in a new global research coalition to identify novel ways to strengthen food safety around the world and improve public health. Participants were invited to attend the upcoming Global Summit on Regulatory Science, which will take place in August 2014 in Montreal, QC. The Summit will focus on enhancing the translation of basic science into regulatory applications.

Current State – Characterization of Microbial Hazards and Regulatory Decision Making

Public Health Agency of Canada: Food-borne Illness Outbreak Investigations from An Epidemiologic Perspective

Diane MacDonald (Public Health Agency of Canada) presented an overview of food-borne illness outbreak investigations from an epidemiologic perspective. Food-borne illness outbreak investigation and response in Canada involves provincial and municipal agencies, the Public Health Agency of Canada (PHAC), Health Canada (HC), the CFIA, the industry as well as international partners. The federal role in outbreak investigations was described. Food illness outbreak investigations include three elements: laboratory investigation, food safety investigation and epidemiological investigations. The epidemiological aspects of outbreak investigation and related challenges were described. Challenges include determining whether

two potential cases are related, as well as constructing a case definition for confirmed, probable or suspect outbreak cases. Clinical criteria, laboratory criteria as well as other elements of investigation serve in the case definition. Outbreak investigations are complex and multidisciplinary, and new molecular typing methods have an impact on contributing additional information to the epidemiological investigation.

Food-borne Disease Real Time Surveillance and Outbreak Response in Canada

Celine Nadon (Public Health Agency of Canada) provided an overview of current procedures for national outbreak detection and response from the PulseNet's perspective. Her presentation described the role and operations of PulseNet – a Canada wide system, designed to provide the ability to consider all potential outbreak cases, in real time, in order to determine whether cases are related or not. All ten provincial public health laboratories, as well as HC and the CFIA participate in PulseNet. A set of primary and secondary tools are used to analyze samples. Pulsed-Field Gel Electrophoresis (PFGE) represents one of the most important primary tools used; PFGE allows scientists to look for whether isolates have a matching PFGE patterns. Considerations in tool selection include the need for balance between sensitivity and specificity of the analysis, as well as the ability to compare test results among various laboratories using a widespread tools and standardized test methods. Data interpretation criteria used to determine whether there is a match or not, as well as the quality of that match are critical to the investigation. Future projects were also presented in the context of a five-year roadmap to PulseNet Canada Genome.

Local and Provincial Food-borne Illness Outbreak Identification and Response

Lance Honish (Alberta Health Services) presented an overview of the Provincial public health response. An overview of local food-borne illness surveillance and of outbreak identification and investigation was presented, as well as the role genomic tools play in this context. At the local level, outbreaks are identified through information sources such as food-borne illness surveillance systems or reported cases of illnesses. Notifiable disease surveillance systems provide valuable information on potential outbreaks. The next steps include analysis and interpretation to determine whether a possible outbreak warrants further investigation, leading to outbreak response when applicable. The question of whether “commercial foods” are implicated is considered to determine the scope of an investigation: local, provincial or national. Case studies illustrating how these principles are applied to actual cases of outbreaks were presented; one of the cases illustrated remained at the local level and two cases demonstrated how cases identified locally ultimately resulted in a national response. Genomic tools were instrumental in identifying suspect cases at the outset, and in determining the source of the outbreaks.

Application of the Weight of Evidence in Responsive Health Risk Assessments during Food-borne Outbreaks

Enrico Buenaventura (Health Canada) presented an overview of the Health Risk Assessment (HRA) process and the “weight of evidence” concept. The decision making framework includes risk assessment, risk management and risk communication. In order to provide confidence in decision making, an evidence-based approach supports the work and decisions being made. Recommendation No. 29 of the 2009 Weatherhill report led to the adoption of a weigh of evidence approach that considers factors for

appropriate and timely action in a food-borne illness outbreak investigation. The health risk assessment process includes an analysis of hazards, exposure assessment and risk characterization (estimation); this is followed by documentation and communication. Microbiology, epidemiology and food safety investigations represent three streams of evidence which are considered in a weight of evidence approach in the health risk assessment. Examples were presented to illustrate how these three streams of evidence are used in actual cases.

Discussion

Following these presentations, workshop participants considered the following questions:

- Based on what we know today, to what degree do we think genomics can contribute to food inspection in Canada?
- What are our key preoccupations in the characterization of bacterial isolates in food-borne disease outbreak investigations today? What are the tensions and key issues?

The following points emerged in response to these questions:

Benefits

- Genomic tools have a great potential to contribute to food inspection in Canada, and are already being used.
- In a legal context (e.g. recalls), genomic tools can provide valuable support for epidemiological investigation and source attribution.

Preoccupations

- Standardization is important, as well as validation of test methods.
- Recognizing that different levels of standards may apply in various cases will be important (i.e. fitness for purpose).
- Genomic surveillance of the “non human” side would also be useful, in order to provide additional context.
- Current tools still require an isolate; considerations need to be given to eliminating or leveraging this requirement.
- There are tensions between the need for rapidity and the need for specific and accurate test results. In an outbreak situation, given the public health risks involved, both are important. Tensions exist between speed, accuracy and cost effectiveness.
- Questions such as how to use genomic tools in situations when information gaps exist (e.g. when there is no sample) also need to be considered.
- It would be useful to use genomic tools to further leverage the rich information from existing databases, and to leverage the work of other groups such as university researchers.
- Focusing efforts and resources will be important, as it is not feasible to investigate all the clusters of potential outbreak cases or clusters of strains identified.

- Interpreting the data and determining what is significant vs. not will be critical (i.e. this is the crux of the question). In the context of new tools, there is a learning curve involved in determining when a cluster of strains is meaningful and requires further investigation.
- There is a need for rapid tests that lead to the development of intervention strategies.
- Developing prevention methods using genomic tools to enhance food quality control in the food industry will also be valuable (e.g. predictive hazard).
- Science often precedes policy development; policies are developed in light of what science has made possible, as well as other considerations.
- Challenges in communicating the information resulting from the use of genomic tools can be expected. This will need to be considered when preparing reports intended for inspectors or decision-makers.

Current Tools for Detection, Typing, Characterization, Risk Profiling – High Risk Pathogens

Implementation of Molecular Techniques in Regulatory Food Microbiology Testing

Burton Blais (Canadian Food Inspection Agency) provided an overview of the implementation of molecular techniques in regulatory food microbiology testing. When testing is conducted to verify compliance with food regulations, detection and typing are important. In outbreak investigations, source attribution, scope of contamination and hazard mitigation are also needed. Timeliness as well as quality assurance and the reliance on validated methods represent important elements in regulatory testing. Phenotypic approaches and genomic approaches to testing were compared. A recent case of *E. Coli* O0157:H7 in ground beef illustrated the benefits of using genomic tools; colonies of food-borne *E. Coli* O0157:H7 were identified on the basis of key marker genes. This approach resulted in a significantly reduced turnaround time for confirmation, more timely regulatory interventions and demonstrated the effective use of a genomic-based method in a food inspection scenario. The identification of priority verotoxigenic *E. coli* (VTEC) was also presented to illustrate how various tools can be used together. Genomic tools applications to food pathogens include identification, virulence profiling and sub-typing, leading to the creation of risk based inspection tools.

From A,C,G,T to PFGE, to MLST, to MLVA, to WGS: How to Make Food Safer

Franco Pagotto (Health Canada) provided an overview of various tools – PFGE, Multilocus Sequence Typing (MLST), Multiple-Locus VNTR Analysis (MLVA) (VNTR: Variable-Number Tandem Repeat) and Whole Genome Sequencing (WGS) – illustrating their similarities and differences. Microbial WGS has several advantages including the fact that it offers a comprehensive, unbiased approach and is now amenable to single cell analysis. Potential challenges with databases were outlined to illustrate the criticality of database construction; bioinformatics challenges are also expected as WGS approaches are implemented. In terms of developing tools and applications, it is important to consider what this means with respect to food safety and what are the regulatory and industry perspectives to consider.

Use of Molecular Markers for Characterization of Food-Borne Pathogens

Cecile Tremblay (Laboratoire de Santé Publique du Québec) presented molecular approaches for the characterization of food-borne pathogens. Challenges in this area relate to the fact that health systems jurisdictions are national and regional, whereas bacteria are ubiquitous; microbes tend to be studied individually; and surveillance systems are fragmented. Genomic tools can provide new avenues for solutions to these challenges. Examples of WGS applications were presented to illustrate the possibility of using these tools effectively. Investigations of cases involving a *Legionella* outbreak, the emergence of a multidrug resistant *Salmonella Dublin* from human and animal sources, as well comparative genomic analysis of two *Vibrio Cholerae* strains related to Haiti travel illustrated the benefits of using WGS. *Salmonella Enteritidis* and *Salmonella Heidelberg* can be analyzed using PFGE, however this technique has limitations and WGS provides the ability to further discriminate between other characteristics, relative to PFGE. WGS will create challenges in the area of data analysis, as the interpretation can be different depending on the algorithm used. Standardization is key to share data and ensure quality control.

Front Line Applications of New Typing Methodologies to Investigate Food-Borne Disease: Promises and Challenges

Vanessa Allen (Public Health Ontario) presented the front line application of new typing methodologies to investigate food-borne disease, as well as how these methodologies present both promises and challenges. Her presentation covered current tools that build on past successes for outbreak detection and response, genomic technologies and their initial applications, as well as considerations to take into account in implementing WGS tools in order to ensure the best public health action – how best to equip front line users with knowledge and expertise they can use to take proper action. The advantages and disadvantages of PFGE and WGS for typing of food-borne bacteria were compared and examples of outbreak investigations were provided. A number of conclusions were drawn from the experiences described. The assessment of protocols for technical aspects of the development, assembly and analysis of genomic data, as well as validation of epidemiological concordance will be important; thus both technical and epidemiological validation are required. Data integration and computational tools open up an array of new possibilities; however data storage is expected to be a challenge. User need assessment and knowledge translation will be critical to the success of public health action based on genomic tools.

Discussion

Participants discussed the following questions:

- To what degree are genomics being used today in both the characterization of microbial hazards and regulatory decision making?
- What have we heard this afternoon that includes a genomics component?

The following points emerged in response to these questions:

Developing and implementing new approaches

- Defining the questions will drive the genomic tools and applications which end up being developed.
- Reviewing lessons learned from the introduction of prior tools such as MLVA would be useful to ensure that these learnings are applied to the introduction of genomic tools.
- Developing a judicious approach to develop and identify readily deployable tools that front line laboratories could use would be valuable (i.e. reference centers laboratories could develop and validate methods that other laboratories could then apply).
- Technologies to isolate or concentrate the sample could be beneficial, given that generating the isolate is a time consuming step.
- A concern about the potential high cost of bioinformatics was expressed.
- Training costs for personnel also need to be considered.

Managing and harmonizing the data

- Managing the amount of data generated will be important, as well as leveraging and sharing existing databases.
- The question of open source software and/or open source data will need to be considered. A potential issue with open source software is that it could be modified between the time a test is run and repeated, potentially leading to different interpretation of results. Security issues are also a concern.
- In terms of harmonizing databases as well as data quality, considerations need to be given to the type and extent of harmonization required. Do databases need to be harmonized at the process level, or is the ability to exchange data among different systems sufficient? (i.e. inter-operability). Similarly for data quality: does the entire process need to be harmonized or should the goal be to achieve comparability?

Interpreting the data

- The question of whether the presence of a gene would suffice to trigger an action in the absence of expression of a toxin will need to be considered (i.e. meaningfulness of a given result). Consideration must also be given to fact that not all toxins may be expressed in Vitro during testing; however, the same toxins could be expressed in Vivo causing illness. In some cases, the presence of toxin may be much better detected and characterized using tests targeting the genes rather than tests targeting the expressed protein.
- Determining validation requirements will be important (e.g. what are the elements that require validation, process, data analysis, algorithm, etc.; the ability to withstand legal scrutiny in the case of legal cases will be required – i.e. requirement for an isolate).
- Standard reference materials or calibration samples can be used.

Developing potential applications

- Generating databases (e.g. proteomics, mass spectral analysis, etc.) could be useful as well as conducting comparability analysis and targeting specific proteins and biomarkers.

- Conducting metagenomics analysis in food samples would provide the ability to mine the information generated, using different markers for various purposes (e.g. prospective and retrospective analysis).
- Developing rapid tools which can be used at the farm or slaughterhouse level would be useful for prevention: pathogens in foods can be controlled most successfully when intervening at this stage.
 - The industry would benefit from the ability to detect an issue and prevent an outbreak before it occurred. However, the industry would likely be reluctant to use such approaches if they result in traceable information.

Communicating with end-users

- Educating the end-users about the possibilities and the limitations of genomic tools will be important, as well as providing assistance with interpretation of the results.

New Possibilities for Genomics

One Disrupting Technology Fits it All: Towards Standardized Bacterial Whole Genome Sequencing for Global Surveillance

Dag Harmsen (University of Muenster), presented an approach using one disruptive technology, leading to standardized bacterial whole genome sequencing for global surveillance. A review of various projects and publications was presented to illustrate various applications and their related strengths and challenges. The importance of developing standard nomenclature was highlighted, as this will be necessary, in order to collaborate internationally and ensure consistency in the exchange of information or data.

The phylogenetic analysis of *EHEC* O104:H4 by hybrid reference mapping and *de novo* assemblies and BIGSdb core genome MLST (cgMLST) illustrated that it is possible to conduct this type of analysis in real time. Other work where substitutions, insertions and deletions were analyzed revealed that it is gene by gene *de novo* consensus accuracy for PGM, MISEq and GSJ that matters. To enhance the specificity of surveillance, four dimensions need to be considered: place, time, person and type. Current bottlenecks include sample processing, NGS platforms, bioinformatics and information technology infrastructure as well as well trained personnel.

In conclusion, a standardized hierarchical microbial typing is proposed. This tiered approach would rely on methods with increased discriminatory power, as required through the different tiers. First, MLST, SNPs, and rMLST would be used; the second tier would rely on MLST+; finally, should additional tests be needed, SNPs/Alleles testing could be conducted.

Identification and Characterization of Food-Borne Pathogens by Whole Genome Sequencing: a Shift in Paradigm

Peter Gerner-Smidt (Centres for Disease Control and Prevention, CDC) provided an overview of how the identification and characterization of food-borne pathogens by WGS is leading to a shift in paradigm. At the CDC, WGS began with the investigation of the Haiti cholera outbreak in 2010. The CDC is conducting a

proof of concept study on the use of real time WGS in conjunction with enhanced surveillance for *Listeriosis*.

In terms of software requirements, it was noted that software for use in public health networks must be comprehensive, run on standard desktops, and be user-friendly. Moreover, the output produced must be simple and easy to communicate and the data must be located in central and local databases. To date, *BioNumerics* (from Applied Maths) is the software that is most ideal in terms of these characteristics.

The CDC currently uses nine different methods for various pathogens; each pathogen has a specific workflow and various timelines (i.e. from less than a day to many months). The intent is to replace all of these methods by WGS, which can be conducted in three days. This approach will result in a number of benefits to public health: greater efficiency of the laboratory workflow; the ability to conduct outbreak detection, investigation and control with precise case definition; and the ability to conduct surveillance of sporadic infections through efficient source attribution analysis of sporadic disease as well as a focus on pathogens of importance to public health.

Genometraker: a Pilot Source Tracking Network of Next Generation Sequencing Desktop Sequencers

Marc Allard (Food and Drug Administration, FDA) presented an overview of the integration of Next Generation Sequencing (NGS) desktop sequencers to build a global genomic network for pathogen traceback and outbreak detection. Regulators can benefit from WGS to identify source information, to conduct surveillance, to perform risk assessment and modeling and to replace traditional bacteriological typing procedures. During outbreak investigations, key questions include determining whether a particular isolate is part of the outbreak, whether it has been observed in the past and whether it matches a clinical isolate. The Genome Traker project was described. Genome Traker is a new pilot network of State and Federal Public Health Laboratories collecting and sharing genomic data from food-borne pathogens. Researchers around the United States (U.S.) will be able to analyze and compare data in real time, speeding up investigations and contamination control. Archived data will become the foundation for other national and international research platforms. A project involving collaboration with CDC on real-time *Listeria* outbreak detection is currently in progress. The objective is to type every clinical and food/environmental isolate of *Listeria monocytogenes* collected in the U.S.

Future plans include partnering with more organizations that have sequencers and/or isolates; in this realm, partnerships with additional States as well as with other countries and organizations are being envisioned. Other partnerships within the FDA and with other organizations have also been established. For example, certified reference materials for genomic sequencing are available from the National Institute of Standards and Technology (NIST). The future vision is for the data to be publicly available, just as weather data is, so that new tools and ideas can be generated from multiple users and contributors.

The Microbial In Silico Typing (MIST) Pipeline – Prospects for Rapid Analysis and Interpretation of Draft WGS Data from Food-Borne Pathogens

Ed Taboada (Public Health Agency of Canada, PHAC) presented prospects for rapid analysis and interpretation of draft WGS data from food-borne pathogens, through the Microbial In Silico Typing (MIST)

pipeline. He observed that some methods are better suited to short-term epidemiology and others are best suited to long-term epidemiology. Short term epidemiology requires a higher level of discriminatory power, whereas long-term epidemiology requires more “stable” genotypes. What constitutes a good WGS-based method to implement in the epidemiology context may turn out to be specific to each microorganism. Genomic data can improve the method development cycle through the development of better markers than in the past as well as genome-based validation prior to deployment. WGS is beginning to be used in outbreak investigations more and more frequently. Strategies to harness WGS towards sporadic outbreak cases (i.e. large-scale integrated surveillance) will be needed. As WGS costs continue to go down, the paradigm related to the identification of the best approach in a given case is changing. A key question to consider is how much effort to invest in transitional typing methods in the context of WGS becoming more prevalent. Another question is how to shift from WGS-informed methods to WGS-based methods.

Genomics Unknots a Rather Knotty Issue – the SNPing of *Salmonella Enteritidis*

Dele Ogunremi (Canadian Food Inspection Agency) presented work conducted to develop typing assays for differentiating *Salmonella Enteritidis*. *Salmonella* is an important food-borne pathogen, with a high burden of illness and a wide distribution. This work illustrates a case where genomic tools helped solve a complex issue, through the SNPing of *S. Enteritidis*. PFGE does not generate useful discrimination between sub-types of interest, creating limitations for surveillance or outbreak investigation. The strategy for developing a subtyping tool for *S. Enteritidis* was outlined. DNA was isolated from pure bacterial isolates, then whole genome analysis was conducted using two different platforms; the genome was assembled (*de novo* and reference assembly), resulting in genome annotation. This work was carried out with 334 isolates; signature SNPs could be derived based on the results. In conclusion, by using genomic tools, an SNP subtyping tool was developed for use in support of food safety investigations and surveillance.

Discussion

Participants discussed the following questions:

- What are the possibilities for genomics to support food-borne disease outbreak investigations?
- What are the limitations or anticipated challenges?

The following points emerged during the discussion:

Possibilities

- Sequencing historical strains and leveraging that information in comparison with PFGE patterns could provide useful comparisons.
- Recognizing that different pathogens will require different approaches is key (i.e. different situations and different pathogens will dictate a variety of approaches).
- Efforts to engage with end-users such as epidemiologists will help develop approaches that meet their needs.

- An analogy with weather and hurricane patterns data was made: identifying meaningful clusters of outbreaks for end-users is akin to genomic scientists (involved in investigating suspected cases of outbreaks) “naming the storm.”

Potential challenges

- Considering how to keep up with standards can be a challenge. Standardization and interpretation will be critical elements to consider, as new tools are developed and implemented.
- Determining the approach in creating or owning databases and/or using common databases (e.g. sharing and storing information) will be important. Data management and information technology issues will need to be considered.
- Budgetary limitations will drive the development of strategies (e.g. collaborating on efforts towards sequencing strains of interest over time, etc.).

The Path Forward

Discussion on Other Challenges and Additional Research Needed

In the final discussion, workshop participants considered the following questions:

- What are the additional challenges for food-borne disease outbreak investigations that current technology may not address?
- What additional research is implied? What are the questions that research could help answer? What type of information needs to be extracted from food isolates to support outbreak investigations?

The following points emerged during the final discussion:

Potential developments and applications

- Knowing what genes are expressed could be useful (i.e. WGS does not provide functional data).
 - Other views were expressed on this point: expression data may not be required for an investigation, and/or could be determined afterwards if there is a research interest.
- Finding ways to reduce the time required to obtain the isolate would be useful, however this is limited by the speed of growth of bacteria. On the other hand, there are increasing ways of conducting analysis with smaller amounts of isolates, which also contribute to reducing the amount of time required before the analysis can be performed.
- Sequencing all the historical strains would be valuable, in terms of future research.
- Applications of WGS are possible for culturable virus and parasites, in addition to bacteria.
- Applications to drinking water could be of interest, as many organisms present in water can cause infections if ingested by a mammalian, but they would not grow on a culture.
- Developing preventative approaches would be very valuable in the future.

- Databases could be built and both prospective and retrospective approaches could be developed. For example, it would be possible to predict the emergence of strains or identify those that have the potential to create an illness.
- Considerable research would be required to develop these proactive and preventative approaches.

Jurisdictional, legal and standardization considerations

- Data release agreement protocols will need to be established (e.g. federal vs. provincial responsibilities in various circumstances depending on which organization generated the data, etc.).
- Standardization will be critical, in order to sustain legal challenges (e.g. source attribution, regulatory actions, etc.).
- Determining and agreeing on criteria for the interpretation of the data will also be crucial.

Costs, technology and human resource considerations

- Given that the cost of sampling and isolating the material is significant; in this context, the proportion of costs attributable to WGS may be minimal (e.g. only 1 % of the food tested is positive and requires further investigation).
- Sustaining bioinformatics requires considering questions such as human resources competency, training and/or access to bioinformatics centers. In the context of recruiting workers for short term projects, this can pose additional challenges.
- Addressing information technology needs (e.g. data storage, access, security, data ownership, centralized vs. local databases, etc.) may be a research area in itself, in terms of determining the best approaches in this area.
 - Regulatory users need to retain the ability to generate and/or interpret their data (e.g. issue if a regulatory action was taken based on someone else's results or interpretation).

A change in paradigm

- It is expected that most barriers will be social, and not technological: resistance to change, communicating limitations and benefits, understanding and collaborating with end-users, etc. Approaches such as investing in capacity building and building partnerships will be valuable to overcome these barriers.
- Regulatory considerations may need to be addressed (e.g. do genomic tools have the potential to affect the definition of a pathogen?).
- The ability to retrospectively understand pathogen changes and to predict emerging pathogens may also affect how food inspection is approached overall (i.e. paradigm shift).

Proposed Next Steps

The following next steps were outlined:

- A workshop report will be prepared and circulated.

Supplemental information. *A concept paper outlining a vision for the integration of genomics in food inspection has been prepared by a sub-group of the Organizing Committee (Appendix “A”). This may serve as a focal point for on-going discussions on best practices for implementation of genomics in regulatory food safety.*

- Feedback will be sought from participants regarding the identification of potential research priorities and towards the development of guidance for laboratory scientists to help ensure that the end-users can best benefit from the information generated.
 - Input from participants on research priorities will be useful for the CFIA, HC and PHAC when evaluating the value of research proposals in light of their outcomes and applications.

Survey conclusions. *At the time of this writing, input has been received through the administration of a survey designed to evaluate the priorities and preoccupations of participants representing different perspectives in the food safety sciences. The results of the survey are presented in Appendix “B”.*

The following indications and priorities emerge from this exercise:

- *Stakeholders at all levels acknowledge the legitimacy of genomics approaches in characterizing food isolates to underscore regulatory decisions*
- *Timeliness in the delivery of analytical results must not come at the expense of accuracy*
- *Harmonization of genomic approaches and databases will be important*
- *Sharing genomic information among all levels of government is important*
- *Quality assurance and standards for best practices are important*
- *Reports of analysis for end-users should be presented in a clear, concise format*

Closing Remarks

Burton Blais thanked the committee members, the presenters as well as many who contributed to the development of science and demonstrated what could be achieved through research and development. He noted that he was very pleased with the diversity of representation of participants, from researchers to end-users. The workshop provided a valuable forum and rich discussions to further the understanding of the reality of both sides.

Appendix A – Concept paper

A Vision for the Integration of Genomics in Food Inspection

Burton Blais*, Catherine Carrillo and Dominic Lambert

Research and Development Section
Ottawa Laboratory Carling
Canadian Food Inspection Agency

(* Contact for further information: Burton.Blais@inspection.gc.ca)

Summary

Recent developments in the field of pathogen genomics herald a new paradigm for analytical food microbiology in which pathogenic bacteria will be characterized on the basis of their genetic profile rather than traditional approaches relying on morphological behaviours. The ability to identify gene markers associated with virulence and other properties relevant to the identification, risk profiling and typing of food-borne bacterial isolates will play a critical role in informing regulatory decisions and tracing sources of food contamination. Here we present several scenarios illustrating prospective roles for pathogen genomics in food inspection.

Background

Regulatory food safety agencies worldwide pursue a scientifically informed, risk-based approach in protecting consumers from preventable illnesses. The application of leading edge analytical technologies for the detection and characterization of food-borne pathogens is one of the underpinnings of an effective risk-based inspection system. Approaches capable of maximizing the amount of information obtained in the course of conducting laboratory testing of inspection samples will foster the most appropriate regulatory responses, for example, by informing the health risk assessment process undertaken to categorize the degree of risk attending a contamination incident.

In the present era of globalization and the introduction of new food manufacturing, distribution and consumption practices, food microbiology testing programs require high throughput analytical technologies providing actionable results within a short time frame. The classic approach has relied on recovery of microorganisms from food samples by enrichment and their identification on the basis of phenotypic characteristics elucidated by biochemical and serological techniques. While effective under certain circumstances, there are shortcomings to such a limited approach, for example, when dealing with novel pathogens or trying to attribute contamination sources.

Leading edge genomics technologies open new possibilities for comprehensive analyses of microbial isolates recovered from inspection samples. Next generation sequencing technologies can now render a bacterial genome much faster (i.e., within a working day) and at a significantly lower cost (about one hundred dollars) than previously possible, making it feasible to sequence food-borne isolates in near real-time under certain circumstances (e.g., during food-borne illness outbreak investigations). Though the assembly of finished genomes remains a time-consuming bottleneck, currently available bioinformatics

tools are sufficiently advanced to enable the rapid processing of raw sequence data into usable form for many purposes. Sequencing pathogenic bacteria, whether in the context of outbreak investigations or information gathering in the course of research, can yield an unprecedented quality of information regarding the presence of virulence and other marker genes of relevance to pathogen identification and risk characterization. Furthermore, the identification of unique molecular signatures will enable testing labs to implement customized tests addressing specific strains of interest in determining the scope of contamination events.

While genomics already plays a significant role in the clinical sciences, its role in food microbiology inspection programs remains to be defined. This essay addresses several possible avenues for the integration of genomics technologies in food inspection. Our purpose is to highlight some of the opportunities for present and near-future implementation, and to promote discussion on the promise and limitations of genomics technologies, the identification of research gaps, and the development of guidelines on best practices for integration in regulatory decision making processes.

Strategies for the integration of genomics in food microbiology testing

Here we present different concepts to illustrate how genomics technologies are currently being used in food inspection and prospects for the short-to-medium term (i.e., next 3-5 years). While some of the scenarios are hypothetical, there is a very good base of knowledge and technological capability underscoring their feasibility. Many of the featured examples are predicated on the verotoxigenic *Escherichia coli* (VTEC), a family of pathogens which has been the subject of much regulatory scrutiny in recent years. The VTEC present a suitable model upon which to exercise these concepts since their relatively recent emergence as a food safety concern bears many features amenable to genomic interpretation.

(1) Identification of pathogenic bacteria on the basis of genomic markers

Concept: *Pathogens recovered from foods are identified on the basis of their definitive genetic characteristics using rapid methods such as PCR technology, enabling delivery of test results days sooner than traditional biochemical techniques. DNA-based methods for the identification of food-borne bacterial isolates support more timely regulatory interventions.*

Case study

Escherichia coli O157:H7 have been implicated in outbreaks of food-borne illness associated with the consumption of contaminated foods such as ground beef. In the event of an outbreak it is imperative that production lots associated with the primary food vehicle are identified as quickly as possible in order to recall all contaminated products from the marketplace. Traditional techniques for the detection of *E. coli* O157:H7 in foods rely on a multi-step process involving pre-enrichment in a selective broth followed by plating to reveal the presence of sorbitol-negative colonies, which are then purified and subjected to a battery of biochemical and serological tests to confirm their identity. This process can take up to one week to complete before the contaminant can be definitively identified because of the requirement for growth and phenotypic expression of the organism.

As an alternative to classic phenotypic techniques, the identification of food-borne colony isolates can be achieved on the basis of detection of defining gene markers. Detection platforms incorporating

polymerase chain reaction (PCR) techniques are particularly well suited for same day analysis of a primary colony isolate. The Canadian Food Inspection Agency (CFIA) microbiology laboratory network has undertaken a program of method development aimed at the rapid identification of colonies isolated on plating media at an early stage during the enrichment process. A key technology platform adopted by CFIA for this purpose is the Cloth-based Hybridization Array System (CHAS) providing for identification of pathogens through amplification of key target genes by multiplex PCR, followed by rapid hybridization of the amplicons with an array of immobilized capture probes on a polyester cloth support. This approach enables facile detection of many gene markers in a single reaction, with specificity assured through the hybridization process.

A CHAS method for the identification of *E. coli* O157:H7 was recently validated following the guidelines of the joint Health Canada-CFIA Microbiological Methods Committee (MMC). This method has been published in the *Compendium of Analytical Methods* (MFLP-22: Characterization of verotoxigenic *Escherichia coli* O157:H7 colonies by polymerase chain reaction and cloth-based hybridization array system, B. Blais *et al.*, 2013), enabling its implementation for regulatory testing purposes in Canada. It is notable as the first instance of a genomics-based approach for the definitive identification of a food-borne pathogen isolate. The *E. coli* O157:H7 CHAS was used by CFIA laboratories on two separate occasions in 2013 to provide critical evidence supporting health risk assessments in connection with food-borne illness outbreaks implicating ground beef distributed in Canada. This method enabled the testing laboratories to issue official results of analysis a full two days ahead of the traditional approach, leading to more timely interventions minimizing public exposure to the contaminated product.

Non-O157 VTEC, particularly strains bearing certain O antigenic determinants, are emerging as a serious food-borne public health concern. Unlike *E. coli* O157:H7, there are no biochemical features by which these so-called priority VTEC strains can be differentiated from commensal *E. coli* or other VTEC which are not a public health concern. However, it is universally recognized that food-borne VTEC posing a public health risk can be defined on the basis of possession of certain gene markers, including the verotoxin genes *vt1* or *vt2*, the intimin-coding gene *eae*, and markers for the specific serogroups of concern. Thus, the priority VTEC constitute a striking example of how genomics technologies can be used to discern the presence of gene markers pinpointing pathogens otherwise not readily amenable to identification by classic means. This in turn has enabled practical strategies for detection of such pathogens during routine monitoring of the food supply to verify industry compliance with food safety regulations. In fact, such an approach is the basis for the Canadian VTEC Method which was developed jointly by CFIA and Health Canada for the detection of this family of pathogens in meats and produce.

(2) Deployment of *ad hoc* methods in support of outbreak investigations

Concept: *Genomic information garnered from clinical bacterial isolates implicated in outbreaks of foodborne illness will be the basis for customization of selective recovery and identification procedures to facilitate their detection in food samples during outbreak investigations.*

Proposed scenario

Despite recent efforts of regulatory food safety agencies to implement test methods targeting defined serogroups of so-called priority VTEC, the history of foodborne disease outbreaks is rife with examples of

causative strains with unexpected characteristics (e.g., the 2011 German outbreak in which the aetiologic agent belonged to serogroup O104, not a designated priority serogroup, and lacked the definitive virulence marker *eae*), making it difficult to anticipate detection methods suiting all contingencies. Detection is further complicated by variability among non-O157 VTEC strains in resistance to selective agents commonly used in enrichment culture techniques, hindering their recovery from foods bearing high levels of background microflora.

The state of the art in next generation sequencing technology is nearing the point where clinical isolates implicated in foodborne disease outbreaks will be routinely sequenced in reference laboratories at an early stage during such events. With the application of appropriate bioinformatics tools to analyze the ensuing data, it should be possible to identify strain-specific marker sequences for the development of customized strain-specific test methods (e.g., PCR tests) that can be rapidly deployed to food testing labs conducting analyses in support of outbreak investigations.

Furthermore, the availability of whole genome sequence (WGS) information for these strains should make it possible to ascertain the presence of traits conferring resistance to antimicrobial agents such as antibiotics, quaternary ammonium compounds, tellurite, etc, suggesting an avenue for the formulation of customized selective enrichment media enabling recovery of specific outbreak strains. This would be a particular advantage in instances where a food matrix (e.g., meats, sprouts, etc.) contains high levels of background microflora, which might otherwise interfere with recovery of the target organism.

(3) Source attribution

Concept: Foodborne pathogens will be typed to a high degree of resolution on the basis of whole genome sequencing. Bioinformatic tools will be used for in silico typing to determine profiles that can be matched to current and historical databases generated through programs such as PulseNet. Statistical analyses can be used to generate a microbial forensic likelihood ratio assessing whether two bacterial DNA profiles match (or cannot be excluded as originating from the same source). Strain-specific signatures will enable prompt attribution of food vehicles in food-borne illness investigations.

Proposed scenario

A. Whole Genome Sequencing for molecular typing

Following detection of a pathogen in foods, (sub-)typing methods are often used to generate a profile of the isolated organism to determine if it has been associated with human illness. Typing multiple isolates recovered from food samples can also provide important information regarding the complexity and source(s) of a given contamination incident. Finally, typing enables tracking of food-borne bacterial strains and is frequently used to support regulatory decisions. While phenotype-based typing methods do not independently provide sufficient detail to support decisive regulatory action, DNA- and PCR-based typing methods have been shown to be valuable tools for subtyping bacterial isolates.

Several molecular subtyping schemes have been developed for each of the priority pathogens (e.g. MLST, MLVA, PFGE, CGF, RAPD). The selection of a typing method depends on a number of factors, including proven utility of the method for the pathogen being investigated. Each method requires costly training of lab personnel and in many cases the purchase of specialized capital equipment. Furthermore, comparisons

of typing data among different strains can only be done in cases where the same method has been applied. In some cases variability in the execution of methods by different analysts or different labs significantly impact the comparability of molecular typing data.

Current typing methods are based on a limited subset of genomic sequences and may lack the discriminatory power to differentiate among organisms. DNA typing profiles from two isolates appearing indistinguishable might be interpreted as evidence that the bacteria have a common source. However, the strength of this type of evidence rests on the extent to which the DNA profile consists of a combination of rare traits. When the traits defining a DNA profile are not rare there is a significant probability that two isolates are in fact unrelated, and that matches are mere chance occurrences. In highly clonal strains (e.g., *Salmonella* Enteritidis), where only a few single nucleotide changes may be observed among distantly related strains, most methods are not sufficiently discriminatory.

WGS provides a high resolution molecular typing platform that can be universally applied to bacterial pathogens. In principle, strains differing by a single nucleotide can be distinguished. Furthermore, WGS can now be done more cheaply than lower resolution methods such as MLST or serotyping and is backwards-compatible with previous methods since typing data can be generated from minimally processed genomic data *in silico*. Strains characterized by WGS can be compared to strains characterized by any other DNA-based subtyping method, enabling optimal use of historical data. Molecular typing data have generally been developed as a surrogate measure of the genetic similarity between bacterial strains. Using databases of WGS information, the utility of existing subtyping methods can be rigorously assessed, and improved subtyping schemes that reflect true strain relationships can be developed.

B. Microbial forensics

Although numerous methods are used by food safety and public health agencies to support regulatory decisions during outbreak investigations, demonstrating that food and clinical isolates originated from the same source remains a major challenge. As the results generated by WGS make their way into situation rooms to guide decision makers, concise metrics for the interpretation and contextualization of genomics-derived data will be required to achieve more precise assessments. The concept of “likelihood ratios” is well known in human forensic sciences where they facilitate the interpretation of DNA profiles in matching individuals to a crime scene. For example, when the DNA profile found on a crime scene matches that of a suspect and there is only a one-in-one million probability that this DNA profile might be found in another individual, there is a strong case linking the suspect to the crime scene.

Food inspectors face a similar situation during outbreak investigations when trying to establish causal links between isolates from different sources. Bacteria may undergo subtle changes in their genomes during the course of a food-borne illness outbreak event, with possible impacts on the typing profiles of clonally related isolates recovered over time. The question arises as to how much change in a genome constitutes a significant difference between individual isolates (i.e., different origins or strains). Through statistical analyses of comprehensive pathogen genome databases it should be possible to develop a likelihood ratio approach to determine the probability of finding a given profile in a defined population, and hence, develop criteria to measure sequence diversity between isolates with different degrees of relatedness, and even among clonally related isolates recovered over the course of an outbreak event. This in turn would provide a greater degree of confidence in attributing the origins of isolates, identifying clusters of food-borne illness, their food vehicles and the scope of contamination. This information can also be used to

revise and adjust detection tools (e.g., PCR primers) to ensure their effectiveness in identifying “moving” genomic targets. The elaboration of a forensic likelihood ratios approach would provide a valuable tool to assess the quality of genomic information underlying regulatory decision making.

C. Attribution of food vehicles through genomic surveillance

The advent of genomic typing augurs well for the creation of highly refined databases of bacterial isolates from various sources (foods, production facilities, farms, environmental and clinical strains) providing high resolution characterization of individual strains with established linkages to their geographic and temporal origins. When supplemented with other extensive databases such as the 100 K genome project, this represents a rich resource from which to draw valuable information linking isolates to their origin in the food production continuum.

With the aid of bioinformatic tools databases can be queried to identify genomic signatures that are over-represented in particular food sources for bacterial isolates. For example, it may be possible to identify specific sequences associated with a given food type, production environment, or country. The ability to discern this type of information would be a tremendous boon for food-borne illness investigations: WGS data could be used to determine the “source signature” of clinical isolates, enabling a highly pro-active approach in rapidly narrowing the field for the attribution of food vehicles. Regulatory food inspection agencies such as the CFIA would have an important role to play in such a scheme. Ongoing, extensive sampling plans will be required to ensure adequate representation of different food production elements, such as food types and geographic provenance. Given that most cases of food-borne illness occur sporadically, this approach would enable public health authorities to track the causes of a larger proportion of cases of food borne illness. This would ultimately lead to a better understanding of foods commonly implicated in disease and the implementation of more effective interventions to reduce the burden of food-borne illness.

(4) Hazard characterization

Concept: *The degree of public health risk posed by a food-borne bacterial isolate may be assessed by determining its risk profile, which is a relative measure of its genomic content of virulence, antimicrobial resistance and epidemiological markers. Additionally, metagenomic data from reservoirs of food microbiota may provide a means of predicting the risk for emergence of novel pathogens in a given food production environment.*

Proposed scenario

A. Analysis of food-borne bacterial isolates

Genomic information is highly complex and there are many knowledge gaps with respect to the significance of various marker genes to public health. Nonetheless, there is a growing body of evidence linking certain well defined gene markers to virulence characteristics of bacteria, for example, the role of intimin (coded by the *eae* gene) in the pathogenesis of VTEC, epidemiological associations between certain serotype features and outbreaks of serious food-borne illness (e.g., *L. monocytogenes* serogroups 1/2a, 1/2b and 4b, VTEC serogroups O26, O45, O103, O111, O121, O145 and O157), and even the type of toxin secreted (e.g., verotoxin 2) and the attendant severity of illness.

In the case of VTEC, regulatory food testing programs currently define priority target strains as bearing markers for verotoxin genes and intimin, in addition to markers associated with a narrow family of O serogroups. However, the question arises whether in the course of conducting routine monitoring of food inspection samples the occurrence of an isolate bearing markers for verotoxin and intimin, but none of the so-called priority serogroups, would be actionable. There are varying subjective opinions on the matter, ranging from a narrow interpretation of test results in which only isolates bearing all of the designated factors are considered hazardous, to the more precautionary approach whereby any isolate bearing both verotoxin and intimin factors, regardless of O serogroup, constitutes a public health risk. There is also evidence suggesting that severity or likelihood of food-borne illness varies with verotoxin type and sub-type (for example, VTEC strains possessing *vt2* tend to be more frequently implicated in cases of severe food-borne illness, whereas *vt2_f* is not associated with human illness), and that this should be a factor in determining the appropriate response to the presence of a food contaminant. Yet another possibility would be to define priority VTEC on the basis of contemporary public health data (reviewed periodically) identifying VTEC serogroups most frequently associated with illness in a given jurisdiction.

Such considerations raise problems for health risk assessment specialists who must interpret laboratory results (among other factors) to determine the degree of risk informing the course of regulatory interventions. It should be possible to devise an objective scheme for rating the degree of hazard associated with a given isolate on the basis of genomic analyses. For instance, the public health and food inspection communities can agree on a list of key factors relevant to the characterization of a given pathogen (Table 1). Since not all factors have the same impact, it should be possible to develop a weighted index approach in which each constituent factor determined by genomic analysis makes up a fraction of a final index value which is proportional to the degree of hazard. Such an index value (or hazard characterization score, HazChar Score) used in conjunction with numerical criteria derived from historical data would be the basis for attributing the degree of hazard associated with a particular isolate, which in turn would enable an objective categorization of risk to inform the appropriate regulatory response.

B. Metagenomic analyses of background microflora.

A modern concept in the study of pathogenic bacteria is the emergence of novel pathogens among commensals in a given environment through the acquisition of virulence factors by horizontal gene transfers from other bacteria. The evolutionary trail of the VTEC family suggests *a priori* transformations of benign *E. coli* strains into virulent VTEC having acquired exogenous DNA segments such as bacteriophage carrying verotoxin genes and pathogenicity islands harbouring host colonization factors. There is evidence that other food pathogens such as *Listeria monocytogenes* strains implicated in serious outbreaks of foodborne illness may have acquired enhanced virulence characteristics through horizontal gene transfer processes. This is believed to occur on a relatively short time scale, perhaps on the order of weeks or months, making the emergence of novel pathogens in food production environments or animal reservoirs in near-real time a significant possibility. Furthermore, food-acquired co-infections may arise in which two or more bacterial strains complement each another, for example, a toxigenic strain lacking adherence factors might colonize a host by cross-utilizing a factor secreted by another strain.

Table 1. Proposed concept for hazard characterization: HazChar Score^a

Key factors	Element	Significance
Primary virulence	Toxin	Presence or absence
	Attachment and colonization	<i>eae</i> , enteroaggregative factors
	Pathogenicity	Pathogenesis mechanisms (e.g., haemolysin)
Severity Modulator	Type	<i>vt1</i> vs. <i>vt2</i>
	Subtype	<i>vt2_a</i> vs. <i>vt2_f</i>
Accessory functions	Antibiotic resistance	Therapeutic
	Antimicrobial resistance	Sanitizers
	Persistence	Biofilm formation capacity
	Pathogenicity Islands	Signatures for novel pathogens
Epidemiological markers	Serotype	Outbreak vs. sporadic vs. nil association
	Phage type	Reservoirs, illness outbreaks
	Molecular type	PFGE cluster
Phylogenetic markers	Genus	<i>Salmonella</i> spp.
	Species	<i>E. coli</i>
	Family or group	VTEC

^a A list of key factors is developed for a given pathogen, and each element is assigned a weighted value based on its significance in human illness. Genomic analysis of a foodborne isolate is conducted using either PCR techniques targeting multiple genes or by whole genome sequencing with the application of bioinformatics tools to determine the presence of targeted features. The individual weighted values are summed giving the HazChar Score, which is then compared against a set of predetermined criteria to categorize the degree of risk.

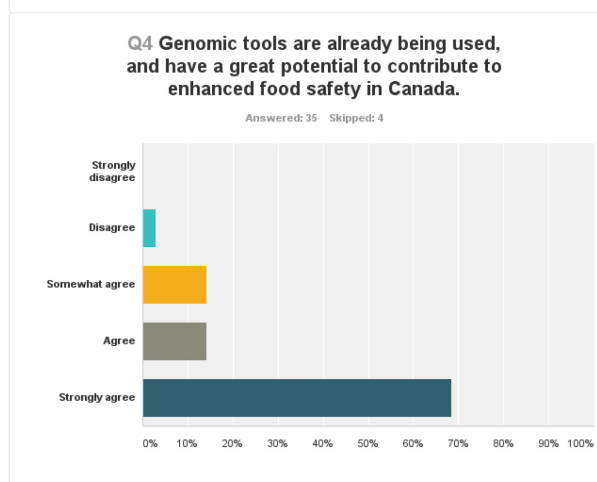
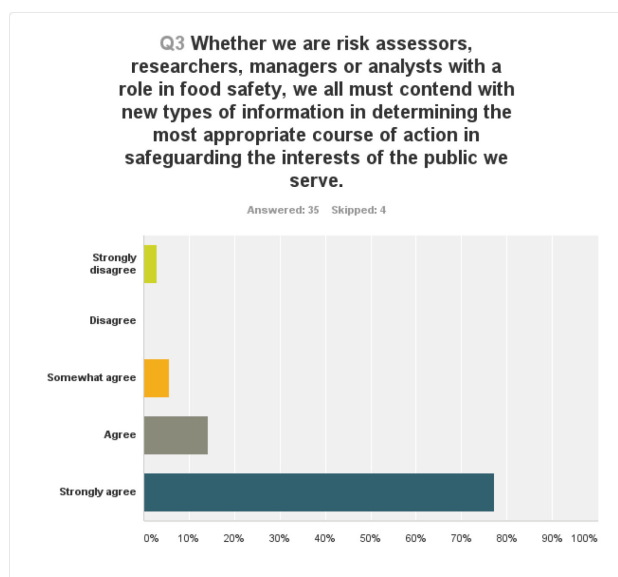
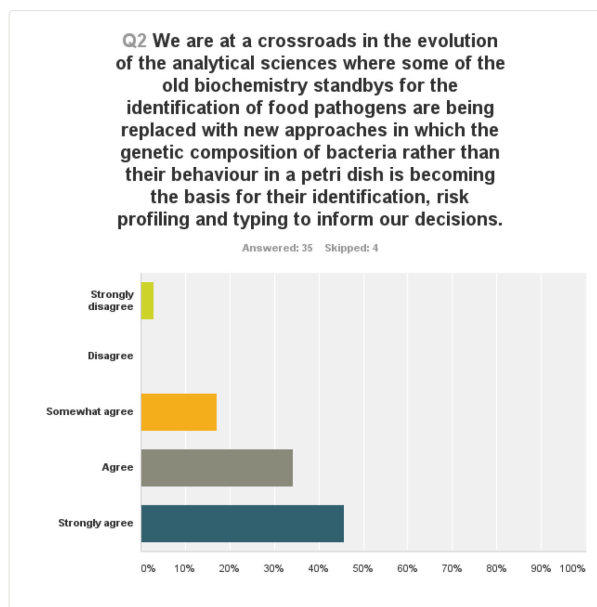
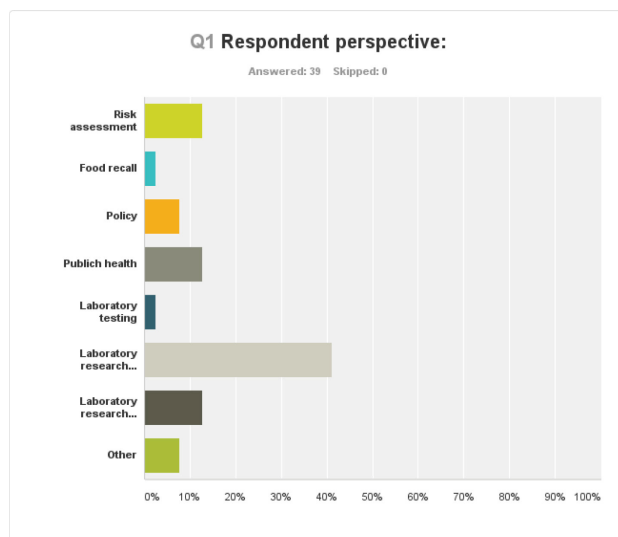
It may be possible to predict the probability of emergence of novel pathogens with enhanced virulence or antibiotic resistance characteristics in the food supply through periodic microbial metagenomic analyses to ascertain the presence of key indicators in the background microflora of food commodities (e.g., ground beef, trim), food manufacturing environments (e.g., floors, food-contact surfaces) and animal reservoirs (e.g., cattle, poultry). A weighted index approach much like that described for the HazChar Score above could be employed here, with possible inclusion of a more comprehensive catalogue of known virulence, AMR and other critical factors relevant to public health.

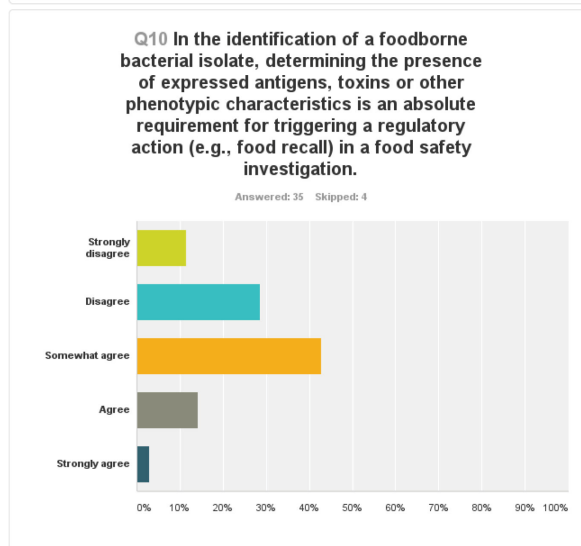
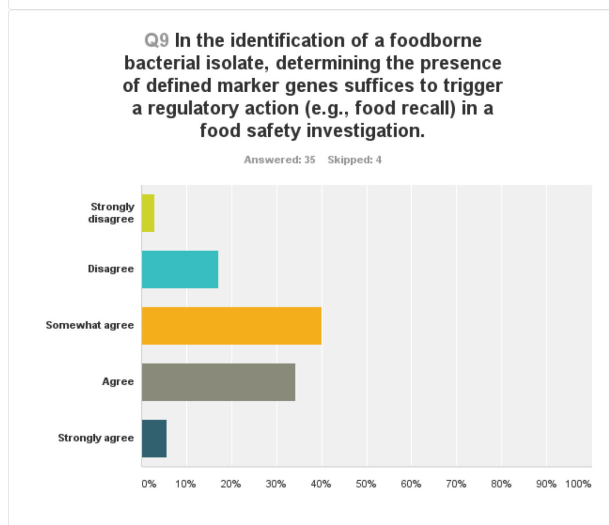
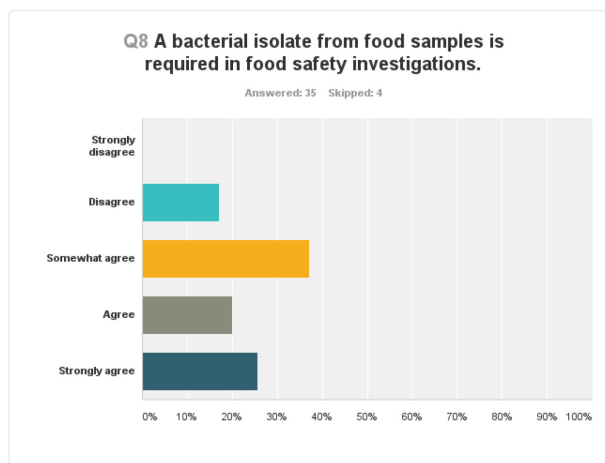
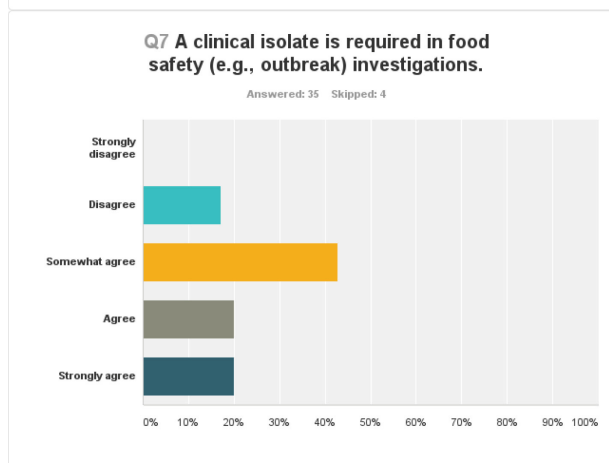
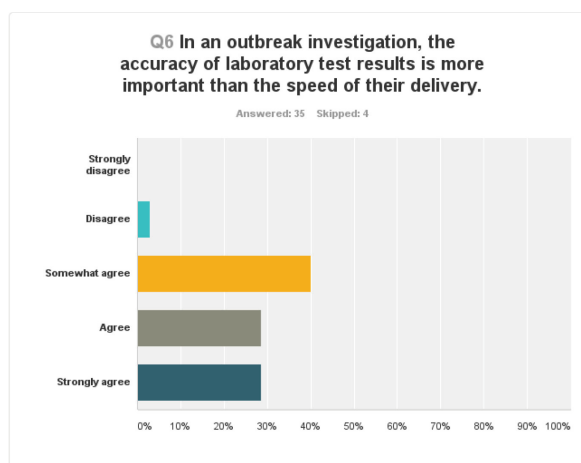
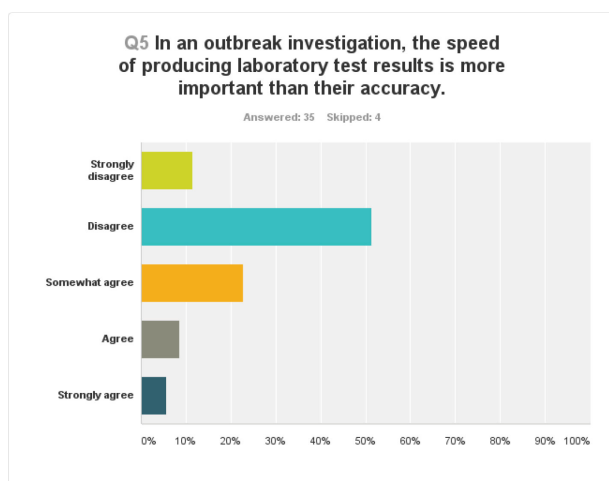
Conclusions

Modern food microbiology research has generated a critical understanding of the epidemiology, pathogenic mechanisms, virulence factors and other salient characteristics of the major food pathogens. The conjunction of expanded scientific knowledge and sophisticated technological capability create exciting new opportunities for the refinement of food microbiology testing programs to meet the needs of a comprehensive risk-based inspection approach. Advances in next generation sequencing technologies have made it possible for investigators to carry out sequencing and processing of bacterial genomes within the time course of a typical food-borne illness outbreak investigation. It may reasonably be expected that

in the near future analysts will be moving from traditional DNA hybridization approaches (e.g., PCR and microarrays) toward rapid whole genome sequencing allowing a much more comprehensive examination of the isolate at hand. This new approach will require access to leading edge bioinformatics capability for analysis of complex genomics data *in silico* to ascertain the presence of key genetic markers (e.g., presence of virulence genes in bacterial pathogens, completeness and functionality of gene products, markers for molecular typing, etc.). The generation and analysis of whole genome sequence information requires the migration of large packets of information between laboratory sites involved in the exploitation of this information, remote computing sites and internet databases for data manipulation and comparative analyses. There are many ways in which the high tech needs of the future can be met, even for relatively small laboratories with low operating budgets. Opportunities abound for food inspection, public health and academic laboratories to pool their resources and serve one another in the common purpose of safeguarding citizens from preventable food-acquired illness.

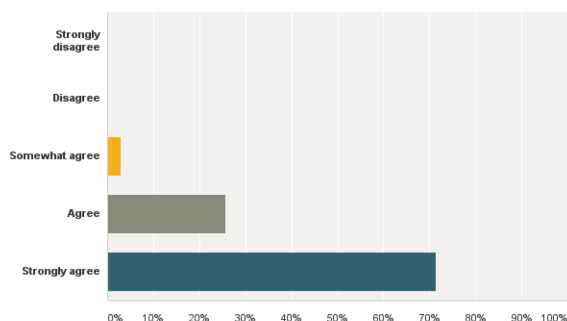
Appendix B – Workshop survey results





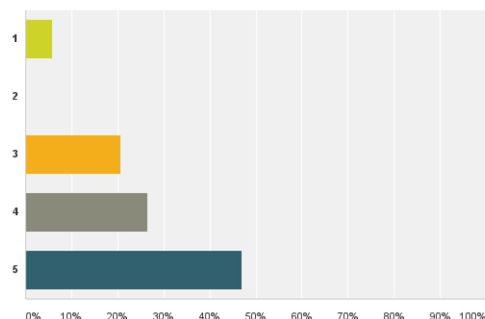
Q11 There is a need for more integration of genomic databases and information sharing at the federal, provincial and international levels.

Answered: 35 Skipped: 4



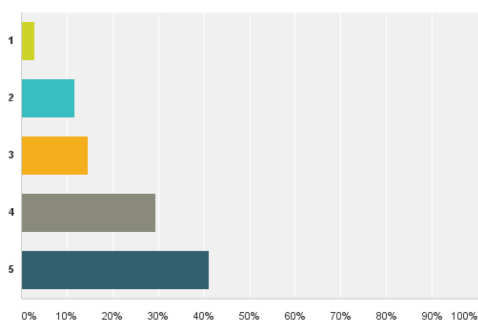
Q12 Validation of genomic methods of analysis to ensure the reliability of data used to inform decision-making.

Answered: 34 Skipped: 5



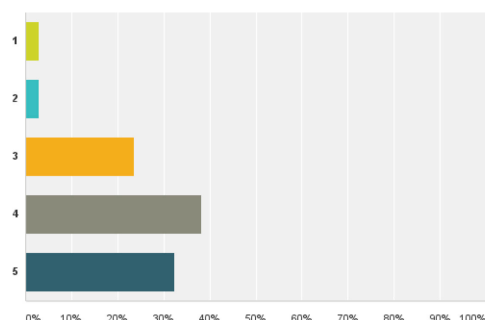
Q13 Adoption of a recognized quality assurance scheme and accreditation of laboratories utilizing genomic methods of analysis that produce results informing regulatory actions.

Answered: 34 Skipped: 5



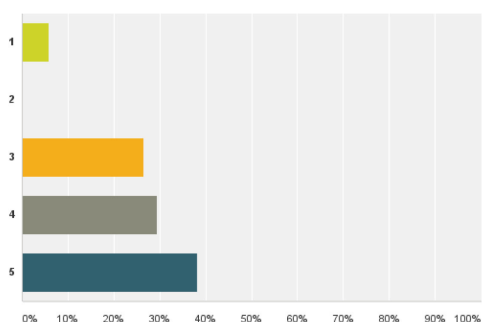
Q14 Performing whole genome sequencing of clinical, food, and environmental isolates on a routine basis.

Answered: 34 Skipped: 5



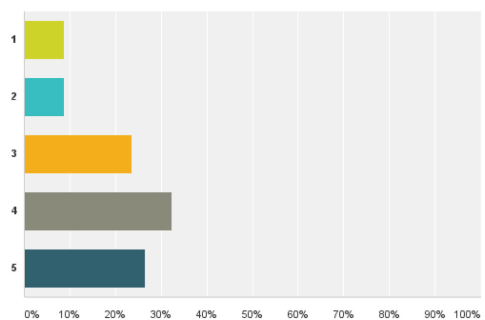
Q15 Harmonizing genomic databases across the food inspection-public health systems.

Answered: 34 Skipped: 5



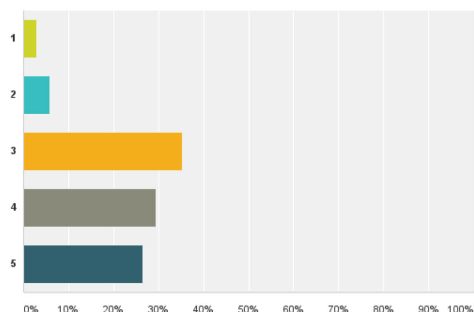
Q16 Ensuring that genomic data and databases are broadly accessible to the regulatory and research communities to encourage innovation and foster best use of available information.

Answered: 34 Skipped: 5



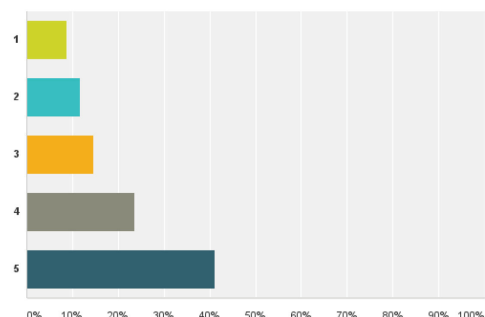
Q17 Finding solutions to deal with sensitive metadata connected with specific bacterial isolates or test samples subjected to genomic analysis.

Answered: 34 Skipped: 5



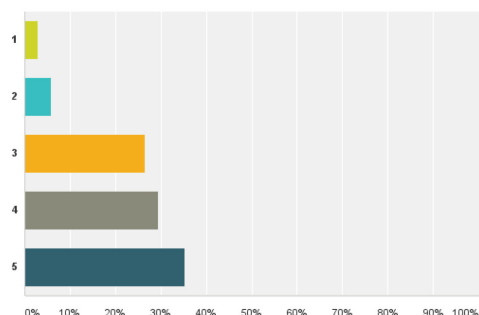
Q18 Development of bioinformatics tools to further leverage the information from existing databases.

Answered: 34 Skipped: 5



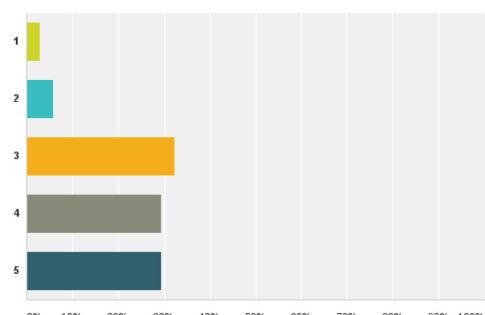
Q19 Development of techniques to expedite isolation and concentration of target bacteria in inspection samples to reduce the turnaround time for delivery of test results.

Answered: 34 Skipped: 5



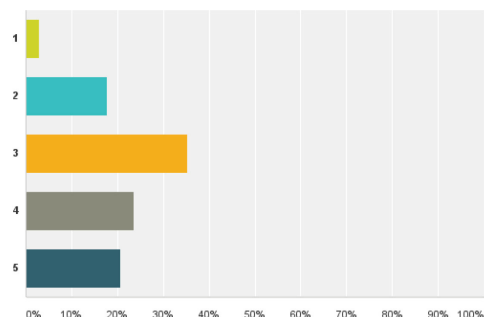
Q20 Development of rapid tests for use at the farm or slaughterhouse level as a means of preventing the entry of pathogens in the food chain.

Answered: 34 Skipped: 5



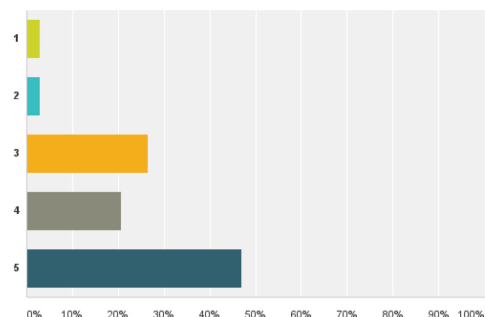
Q21 Development of genomic tools to forecast the emergence of novel pathogens in animal and environmental reservoirs supporting preventive food safety management strategies.

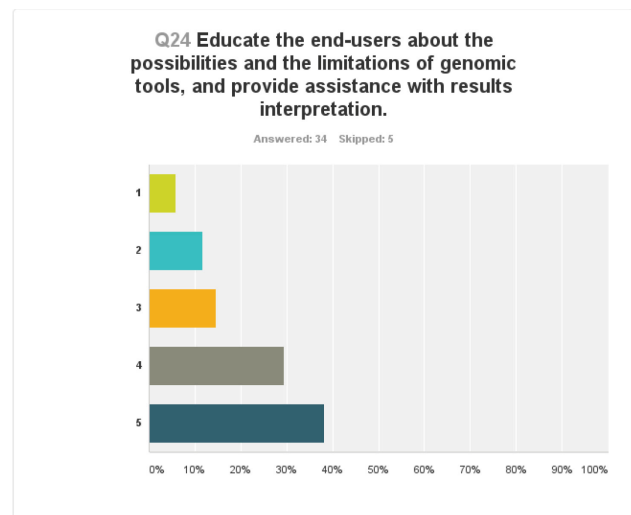
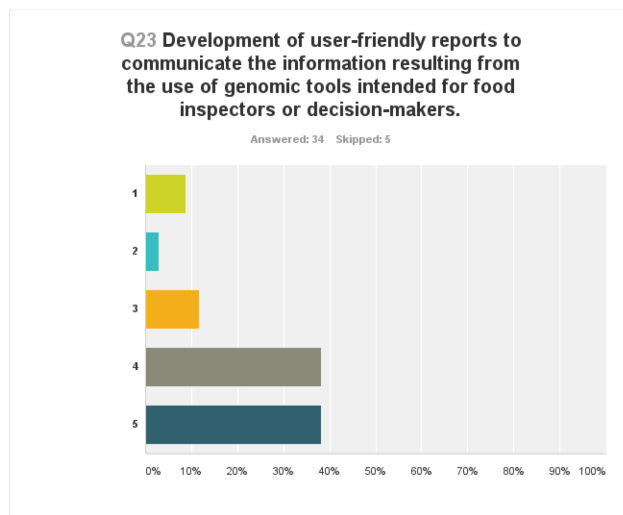
Answered: 34 Skipped: 5



Q22 Conducting genomic analyses of outbreak strains for deployment of specific tests (e.g., PCR) to front line testing laboratories supporting outbreak investigations.

Answered: 34 Skipped: 5





Do you have any additional comments or advice to convey to the Workshop Organizers (these may be integrated in the final workshop report)?

It was an excellent workshop in terms of audience makeup, quality of speakers, facilitation, and outcomes. I'm hoping there will be future workshops on this same issue, as technologies evolve/become more widely used.

These tools should be used to further develop knowledge and capacity in emerging fields such as food virology + research is needed to identify some markers or characteristics that would give us information on whether the organism is alive and infectious, so that we would not need to grow the organisms to prove there is a potential for infection.

1. Absent from this survey and from the meeting was the presence of epidemiologists. The utility of genomics is lost if we cannot engage the epidemiologists, during the uptake of genomics.

2. The use of genomics, beyond a research tool/proof of concept, still needs a lot of quality assurance and quality control work, but in terms of laboratory QC but also utility for epidemiologists.

This was an excellent workshop as it brought together the numerous disciplines that generate and use microbiology and genomics in microbial food safety. While WGS is seen by some proponents as the imminent panacea for all issues in this field, such opinions must be considered with awareness of the need for standardization and validation of the methods to be used in generating, interpreting and reporting the information. This is especially so for regulatory and reference bodies who have provincial, national and global responsibilities.

Great workshop, it would be great if the presentations (ppt) were shared with the participants to facilitate further discussions.

Very glad to have been a part of it, seems like a lot of progress and clarification occurred during the course of the workshop.

Speed over accuracy is a critical factor early in identifying an outbreak (genus, species, +/- serotype). Later accuracy may become more important in identifying cases beyond the index case/isolate (genotype, phylogenetic relatedness). However, even in the later stages (eg XL foods) there may be a very strong pressure to test and declare pos/neg on lots at the same speed as the production line.

What is the legal precedent for demonstration of risk and enforcement of a recall decision (demonstration of an isolate vs PCR)? This will determine what value an isolate has in any investigation.

What role will the regional microbiology labs play in the genomics pipeline? As suppliers of isolates? or as suppliers of sequence data? Remember that couriers do not function on weekends and getting the index isolate to a central sequencing lab will involve delays measured in days.

Do this every year.

Appendix C – Meeting Agenda

Workshop on the application of genomics tools for the rapid molecular characterization of bacterial isolates in foodborne disease outbreak investigations February 24- 25, 2014 – Government Conference Centre, Ottawa

Day 1, February 24

07:30 ‘ Arrival, registration and coffee

A. Getting Started

08:30 ‘ Opening Remarks

- Chair’s welcome - Burton Blais, Canadian Food Inspection Agency
- A vision of genomics for CFIA - Martine Dubuc, Chief Food Safety Officer for Canada

09:00 ‘ Agenda Review Facilitator: Warren Wilson

- Expectations & housekeeping
- Introductions – at tables

B. Current State – Characterization of Microbial Hazards and Regulatory Decision Making

Session Chair: Dominic Lambert (Canadian Food Inspection Agency)

09:15 ‘ Presenters

- Celine Nadon (Public Health Agency of Canada; current procedures for national outbreak detection and response from the PulseNet perspective)
- Diane MacDonald (Public Health Agency of Canada; foodborne illness outbreak investigations from an epidemiologic perspective.)

10:15 ‘ Health Break

10:30 ‘ Presenters

- Lance Honish (Alberta Health Services; Provincial public health response)
- Enrico Buenaventura (Health Canada; HRA process and the “weight of evidence” concept)

11:30 ‘ Discussion # 1

1. Based on what we know today, to what degree do we think genomics can contribute to food inspection in Canada?
2. What are our key preoccupations in the characterization of bacterial isolates in foodborne disease outbreak investigations today? What are the tensions and key issues?

12:00 ‘ Lunch (Not Provided)

C. Current Tools for Detection, Typing, Characterization, Risk Profiling – High Risk Pathogens

Session Chair: Dele Ogunremi (Canadian Food Inspection Agency)

13:30 ‘ Presenters

- Burton Blais (Canadian Food Inspection Agency, Implementation of molecular techniques in regulatory food micro testing)
- Franco Pagotto (Health Canada; "...from A,C,G,T to PFGE...to MLST...to MLVA...to WGS....?")
- Cecile Tremblay (Laboratoire de Santé Publique du Québec; Molecular approaches for characterization of food-borne pathogens)
- Vanessa Allen (Public Health Ontario– Front line application of new typing methodologies to investigate foodborne disease: promises and challenges)

15:30 ‘ Health Break

15:45 ‘ Discussion # 2

1. To what degree are genomics being used today in both the characterization of microbial hazards and regulatory decision making?
2. What have we heard this afternoon that includes a genomics component?

16:15 ‘ Wrapping Up Day 1

- Evaluate the day - Warren Wilson

16:30 ‘ End Day 1

DAY 2, February 25

08:00 ‘ Arrival and coffee

08:30 ‘ Welcome

- Re-cap of messages from Day 1 – Burton Blais/Warren Wilson

D. New Possibilities for Genomics

Session Chair: Catherine Carrillo (Canadian Food Inspection Agency)

08:45 ‘ Presenters

- Dag Harmsen (University of Muenster; One Disrupting Technology Fits it all – Towards Standardized Bacterial Whole Genome Sequencing for Global Surveillance)
- Peter Gerner-Smidt (Centres for Disease Control and Prevention; Identification and characterization of foodborne pathogens by whole genome sequencing: A shift in paradigm)

- Marc Allard (Food and Drug Administration; Integration of NGS Desktop Sequencers to Build a Global Genomic Network for Pathogen Traceback and Outbreak Detection: Description of international (GMI, WHO) and national (GenomeTrakr, 100K) activities.
- 10:15 ‘ Health Break
- 10:45 ‘ Presenter
- Ed Taboada (Public Health Agency of Canada; The Microbial In Silico Typing (MIST) pipeline: prospects for rapid analysis and interpretation of draft WGS data from food-borne pathogens)
- 11:15 ‘ Discussion # 3
- Considering some of the issues and pre-occupations that we discussed on Day 1, and the examples we just heard...
1. What are the possibilities for genomics to support foodborne disease outbreak investigations?
 2. What are the limitations?
- 12:00 ‘ Lunch (Not Provided)
- 13:30 ‘ Presenter
- Dele Ogunremi (Canadian Food Inspection Agency; Genomics unknots a rather knotty issue - The SNPing of *Salmonella Enteritidis*)

E. The Path Forward

- 14:00 ‘ Discussion # 4
3. What are the additional challenges for foodborne disease outbreak investigations that current technology may not address?
 4. What additional research is implied? What are the questions that research could help answer? What type of information needs to be extracted from food isolates to support outbreak investigations?
- 15:00 ‘ Health Break
- 15:15 ‘ Proposed Next Steps - Burton Blais
- The specific process going forward
 - What happens to the outcomes of the workshop?
 - Any feedback or ideas on the proposed path forward?
- 16:00 ‘ Wrapping Up the Session - Burton Blais/Warren Wilson
- Closing Comments
 - Evaluate the workshop
- 16:30 ‘ Adjourn

Appendix D – Participant’s List

Name	Organization
Sara Agha	CFIA
Ray Allain	CFIA
Marc Allard	FDA
Vanessa Allen	OAHPP
Kingsley Amoako	CFIA
Olga Andrievskaia	CFIA
Greg Appleyard	CFIA
John Austin	Health Canada
Micheline Ayoub	Genome Quebec
David Bailey	Genome Alberta
Carole Beaudry	University of Ottawa
Cindy Bell	Genome Canada
Anatoly Belvo	Carlton University
Sabah Bidawid	Health Canada
Luc Bissonnette	Université Laval
Burton Blais	CFIA
Jeanine Boulter-Bitzer	OMAFRA
Enrico Buenaventura	Health Canada
Louise Carriere	CFIA
Catherine Carrillo	CFIA
Angela Catford	Health Canada
Shu Chen	University of Guelph
Kwamaa Duah	CFIA
Nelly Denis	CFIA
Karen Dewar	Genome Canada
Ken Dewar	McGill University
Martine Dubuc	CFIA
Andrée Ann Dupras	CFIA
John Fairbrother	Université de Montréal
Jeffrey Farber	Health Canada
Vic Gannon	PHAC
Colette Gaulin	MAPAQ
Rafael Garduno	CFIA
Peter Gerner-Smidt	CDC
Donna Glezakos	CFIA

Ashkan Golshani	Carlton University
Lawrence Goodridge	McGill University
Dag Harmsen	University of Muenster
Jennifer Holtzman	Health Canada
Lance Honish	Health Services - Alberta
Hongsheng Huang	CFIA
George Huszczyński	CFIA
Irene Iugovaz	Health Canada
Fred Jamieson	CFIA
Karen Jessett	CFIA
Roger Johnson	PHAC
Penelope Kirsch	CFIA
Adam Koziol	CFIA
Michael Knowles	CFIA
Dominic Lambert	CFIA
Franz Lang	Université de Montréal
Linda LeBlanc	CFIA
Min Lin	CFIA
Annie Locas	CFIA
Oliver Lung	CFIA
Diane Macdonald	PHAC
Kim Macdonald	PHAC
Bashir Manji	CFIA
Austin Markell	CFIA
Imelda Marquez	CFIA
Amalia Martinez	Health Canada
Sarah McIlwham	Health Canada
Sam Mohajer	CFIA
John Moisey	CFIA
Susan Nadin-Davis	CFIA
Celine Nadon	PHAC
Dele Ogunremi	CFIA
Katie Omid	CFIA
Denise Oudit	Health Canada
Franco Pagotto	Health Canada
Beverley Phipps-Todd	CFIA
Daniel Plante	Health Canada
Natalie Prystajec	BCCDC
Sean Quinlan	CFIA

Catherine Semple	CFIA
Natisha Stashko	Agri-Food - Alberta
Ed Taboada	PHAC
Monica To	CFIA
Weida Tong	FDA
Tracy Townsend	PHAC
Cécile Tremblay	LSPQ
Linda Vrbova	PHAC
Lisa Waddington	CFIA

Appendix E – Workshop Organizing Committee

Burton Blais, CFIA (Chair)
Catherine Carrillo, CFIA
Dominic Lambert, CFIA
George Huszczyński, CFIA
Donna Glezakos, CFIA
Kingsley Amoako, CFIA
Dele Ogunremi, CFIA
Sam Mohajer, CFIA
Sabah Bidawid, Health Canada
Maria Nazarowec-White, Health Canada

Appendix F – Acknowledgements

The Organizers gratefully acknowledge the support of this initiative by the Canadian Safety and Security Program of Defence Research and Development Canada and the Canadian Food Inspection Agency. We also thank the Intersol Group for facilitating the workshop, note taking and preparing the first draft of this report.